Problem 1: Hidden Markov Model (10pt)

We denote by $\lambda = (N, M, \pi, a, b)$ the Hidden Markov Model, where

- $N$: the number of states
- $M$: the number of possible observations (or tokens)
- $\pi = (\pi_1, \pi_2, \ldots, \pi_N)$: the initial state probabilities, i.e., $\Pr(q_1 = s_i) = \pi_i$.
- $a = [a_{ij}]_{N \times N}$: transition probabilities, i.e., $\Pr(q_{t+1} = s_j | q_t = s_i) = a_{ij}$
- $b = [b_i(k)]_{N \times M}$: token emission probability, i.e., $b_i(k) = \Pr(O_t = k | q_t = s_i)$.

In this homework, you are asked to present an efficient algorithm for computing $\Pr(q_t = s_i | O_1, O_2, \ldots, O_T)$ for $t \leq T$. You are required to submit the details of your algorithm, as well as its computational complexities.

Problem 2: Mixture Models (20pt)

Consider that somebody has a die and a coin. Each time, he can choose to roll the die or to flip the coin. The decision is based on a Bernoulli distribution. The observed number for flipping the coin is either 0 or 1, and the observed number for rolling the die is either 0, 1, 2, 3, 4, or 5. By repeating this experiment for $m$ times, you will observed a sequence of numbers $(n_1, n_2, \ldots, n_m)$. The observation data can be found in http://www.cse.msu.edu/~cse847/assignments/mixture_data.txt.

- Assume that each experiment is independent from the previous one, and meanwhile both the coin and the die are fair with equal probability for each number. Compute the probability $p$ in the Bernoulli distribution using maximum likelihood estimation that fits the observations.

- Still assume that each experiment is independent from others. But assume that we only know that the die is fair one but not the coin. Re-compute your answer for the probability $p$ in Bernoulli distribution, and in the meantime estimate the bias of the coin $b$, which is the probability for observing 1 when randomly flipping the coin.